Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2014, Article ID 898705, 10 pages
http://dx.doi.org/10.1155/2014/898705

Research Article

Building Recognition on Subregion's Multiscale Gist Feature Extraction and Corresponding Columns Information Based Dimensionality Reduction

Bin Li,[1] Wei Pang,[2] Yuhao Liu,[1] Xiangchun Yu,[1] Anan Du,[1] Yecheng Zhang,[1] and Zhezhou Yu[1]

[1] *College of Computer Science and Technology, Jilin University, Changchun 130012, China*
[2] *School of Natural and Computing Sciences, University of Aberdeen, Aberdeen AB24 3UE, UK*

Correspondence should be addressed to Zhezhou Yu; yuzz@jlu.edu.cn

Received 24 October 2013; Accepted 3 April 2014; Published 27 April 2014

Academic Editor: Huijun Gao

Copyright © 2014 Bin Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we proposed a new building recognition method named subregion's multiscale gist feature (SM-gist) extraction and corresponding columns information based dimensionality reduction (CCI-DR). Our proposed building recognition method is presented as a two-stage model: in the first stage, a building image is divided into 4×5 subregions, and gist vectors are extracted from these regions individually. Then, we combine these gist vectors into a matrix with relatively high dimensions. In the second stage, we proposed CCI-DR to project the high dimensional manifold matrix to low dimensional subspace. Compared with the previous building recognition method the advantages of our proposed method are that (1) gist features extracted by SM-gist have the ability to adapt to nonuniform illumination and that (2) CCI-DR can address the limitation of traditional dimensionality reduction methods, which convert gist matrices into vectors and thus mix the corresponding gist vectors from different feature maps. Our building recognition method is evaluated on the Sheffield buildings database, and experiments show that our method can achieve satisfactory performance.

1. Introduction

Building recognition has become an important research area in the field of computer vision. It can be applied to many real world problems, such as robot vision, localization [1], architectural design, and mobile device navigation [2, 3]. Building recognition is a challenging task because building images may be taken under various conditions. For instance, the same building may be taken for images from different viewpoints or under different lighting conditions, and some of these building images may be suffered from occlusions.

Li and Shapiro [4] applied SIFT descriptors to extracting scale invariant features from images. This makes it possible to produce robust and trusted matching among different views of the same building. In [3] the Harris corner detector was employed so that important points for matching buildings can be extracted from the world space for mobile devices. In [5] a hierarchical building recognition approach, which is a two-stage model, was proposed. The first stage of this approach is based on localized color histograms. As for the second stage, SIFT descriptors for local image regions were matched to achieve better recognition results.

As Li and Allinson [6] pointed out, all the above mentioned building recognition algorithms have the following two limitations: first, all these algorithms are based on low-level visual feature detection and the low-level visual features include line segments and vanishing points. So the methods presented in [1–5] are called low-level visual feature methods. The representational performance of these methods is restricted because by purely using low-level features the deeper semantic concepts cannot be better or even possibly revealed. Second, these low-level visual feature methods tend to have very high computational cost and memory demand. This is because the recognition is carried out based on pairs of raw feature vectors, whose dimensions may be very high.

To address the above two limitations of low-level visual feature methods, Li and Allinson [6] developed a new building recognition method, in which the saliency and gist model proposed by Siagian and Itti [7] was used to extract gist features. Siagian and Itti's model was originally used in the task of scene classification. In Siagian and Itti's model 34 feature subchannels are extracted from the original image. Then each feature subchannel is divided into several 4×4 grid subregions and the mean of each grid is taken to produce 16 values for the gist vector. Therefore, the original image can be represented by a feature vector with 544 dimensions. In Li and Allinson's method, to improve the computational efficiency and at the same time preserve as much as possible the useful information for recognition, several dimensionality reduction methods, including principal components analysis (PCA) [8], locality preserving projection (LPP) [9], and linear discriminant analysis (LDA) [10], are employed for dimensionality reduction before classification.

Although the building recognition scheme proposed by Li and Allinson has been proven to be more effective than those low-level visual feature methods [1–5], this method has two limitations: (1) Siagian and Itti's model was directly applied in building recognition, but building recognition is different from scene classification, because building recognition may encounter the problem of nonuniform illumination. (2) The gist vectors extracted from different feature maps were combined to form a 544-dimensional feature vector. This representation may ignore some features which are useful to discriminate the original image from others.

To address the above mentioned limitations of Li and Allinson's method, we propose a novel building recognition method named subregion's multiscale gist feature (SM-gist) extraction and corresponding columns information based dimensionality reduction (CCI-DR), which forms the main contribution of this research.

To address the first limitation, we proposed a novel gist feature extraction method—subregion's multiscale gist feature (SM-gist) extraction. Different from Siagian and Itti's gist model, which builds a Gaussian pyramid of nine spatial scales for each feature channel to extract early visual features from the original image directly, the SM-gist first divides a building image into 4×5 subregions and extracts early visual features from these subregions individually. The reason for doing this is that the lighting conditions of an image may be complicated; however, it is also the fact that the illumination is more uniform in subareas. The Gaussian pyramids built by SM-gist for subareas have five spatial scales. In addition, the center-surround operation is redefined in SM-gist. This is because the size of each subregion is much smaller than that of the original image; if we still create Gaussian pyramids of nine spatial scales, the images in the top level may not contain effective information.

To address the second limitation, we propose the corresponding columns information based dimensionality reduction (CCI-DR). CCI-DR belongs to the graph embedding framework. Different from the state-of-the-art graph based dimensionality reduction methods, which transform gist matrices into their vectorial forms, CCI-DR is based on feature matrices' corresponding columns information, namely, the corresponding gist feature vectors, to determine whether two feature matrices can be neighbors when constructing neighbor graph. By doing this, gist vectors corresponding to different features will not be mixed together, and the features across different building images can be compared.

Many real world datasets of interest are often in high dimension, which brings great challenges in data processing and analysis. So dimensionality reduction algorithm has broad application prospects. Our proposed dimensionality reduction algorithm CCI-DR has great potential to be applied to a wide range of applications, including face recognition [9, 11–14], data-driven fault diagnosis [15–18], switched systems [19, 20], and object category and handwritten digits recognition [21–23].

The rest of this paper is organized as follows. In Section 2 we propose the SM-gist extraction method. In Section 3 we first present the CCI-DR method and then apply CCI-DR to a state-of-the-art unsupervised dimensionality reduction algorithm: locality preserving projection (LPP) [9]. This is followed by the report of experiments in Section 4, which evaluates the performance of SM-gist and CCI-DR. Finally Section 5 concludes the paper.

2. Subregion's Multiscale Gist Feature (SM-Gist) Extraction

It has been suggested by psychophysics research that human has the ability to grasp the holistic information from a scene over a single glance, which is known as the gist of a scene [24]. Siagian and Itti [7] proposed a gist feature and visual attention scene classification framework. Li and Allinson [6] applied Siagian and Itti's model in building recognition. In this section, we will introduce in detail our subregion's multiscale gist feature (SM-gist) extraction.

2.1. The Effect of Nonuniform Illumination on Saliency. Building recognition is different from scene classification, because it may encounter nonuniform illumination. The variety of lighting conditions of an image may be complicated; however, it is also the fact that the illumination is more uniform in local areas. When an image was irradiated by unilateral light, the saliency in some part of the image subject to stronger intensity of light is lower than that in the part subject to the weaker intensity of light. This is because an increase in the light intensity in the background of the image causes the decrease of relative discrepancy in this area compared to the rest of the image. This leads to a big homogeneous area in the image. From information theory we know that the higher frequency of appearance the pixels have, the smaller the amount of information we can obtain. This is the reason why the side of the image containing a bigger proportion of homogeneous area has smaller saliency.

Itti et al. [25] proposed a saliency-based visual attention model, which can extract early visual features. Then these visual features are combined together to form a single topographical saliency map. We utilize the visual attention model to compute saliency map of two images selected from the Sheffield buildings database.

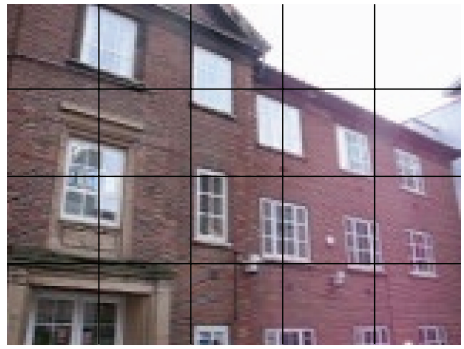


Figure 1: Four-by-five subregions.

Figure 2 shows the effect of nonuniform illumination on the saliency map. Figure 2(a) shows the original map; we can see the light on the right side is stronger. Figure 2(b) is the saliency map of the original map, and we can see that the details of the right side are blurred. Then, we divide the building image into 4×5 subregions (as shown in Figure 1) and compute saliency for each subregion individually. Figure 2(c) shows the saliency map based on subregions. We can see that the details of the right side are well manifested. As Figure 2(d) shows, there is a bunch of sunshine on the right side of the image. The right side of the saliency map calculated by Itti's model is severely blurred, as shown in Figure 2(e), but the right side of subregions based saliency map in Figure 2(f) contains a bit more details.

2.2. Early Visual Feature Extraction. To reduce the nonuniform illumination effects, we propose the subregion's multiscale gist feature (SM-gist) extraction method. A building image is first divided into 4×5 subregions and early visual features are extracted from these regions individually. For each subregion, we extract early visual features, including color, intensity, and orientation, and these features are extracted in parallel. Figure 3 shows the progress of visual features extraction.

As in [25], the intensity of each subregion is computed by the following equation:

$$I = (r + g + b). \tag{1}$$

We obtain four color channels R (red), G (green), B (blue), and Y (yellow) [25] according to

$$\begin{aligned} R &= r - \frac{(g+b)}{2} \\ G &= g - \frac{(r+b)}{2} \\ B &= b - \frac{(r+g)}{2} \\ Y &= \frac{(r+g)}{2} - \frac{|r-g|}{2} - b. \end{aligned} \tag{2}$$

In the above r, g, and b are red, green, and blue channels of the RGB color space of the original image, respectively.

Because the size of each subregion is much smaller than the original image, for each subregion's intensity I a dyadic Gaussian pyramid $I(\sigma)$ of five spatial scales is created by a linear filter, where $\sigma = 1, \ldots, 5$. Similarly, for each color channel four Gaussian pyramids $R(\sigma)$, $B(\sigma)$, $Y(\sigma)$, and $G(\sigma)$ are created. Different from Itti's model our center-surround operation is redefined as follows: the center is a pixel at scale $c = \{1, 2\}$ and the surround is the corresponding pixels at scale $s = c + \delta$, where $\delta = \{2, 3\}$. We obtain feature maps through applying the center-surround operation to Gaussian pyramids. We get four intensity contrast feature maps by (3) as in [25], where $\ominus$ denotes the across-scale difference between two images in a Gaussian pyramid. Then, eight color contrast images are obtained by color pairs red-green as shown in (4) [25] and blue-yellow as shown in (5) [25]:

$$I(c, s) = |I(c) \ominus I(s)|, \tag{3}$$

$$RG(c, s) = |(R(c) - G(c)) \ominus (G(s) - R(s))|, \tag{4}$$

$$B(c, s) = |(B(c) - Y(c)) \ominus (Y(s) - B(s))|. \tag{5}$$

For each subregion of the original image, we use Gabor filters [26] with 4 different scales $c = \{1, 2, 3, 4\}$ and 4 different angles $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ to extract 16 orientation feature maps [25].

After features of all the subregions have been extracted, we assemble the subregion feature maps which represent the same early visual feature into an assembled feature map according to the relative positions of these subregion feature maps to the original image. By doing this, we will obtain 28 assembled feature maps for each building image, and each assembled feature map contains 4×5 subregions. Each assembled feature map represents one particular feature of the original building image (e.g., color, orientation, or intensity).

2.3. Gist Feature Extraction. A gist feature is obtained from an assembled feature map by taking the mean of each subregion to produce 20 values for a gist column. In total, 28 gist columns are computed: 4 for intensity, 8 for color, and 16 for orientations. Then, all the gist feature columns computed from the 28 assembled feature maps are combined into a matrix, which we named the gist feature matrix. The gist feature matrix will be utilized to describe the original building image. Figure 4 shows this process.

3. Corresponding Columns Information Based Dimensionality Reduction

A building image is represented by a gist feature matrix with dimension being 20×28. The gist feature matrix was introduced in the previous section and can be seen as a low-dimensional manifold embedded in a high-dimensional space. Consequently, dimensional reduction algorithms must be applied to gist feature matrices for building images.

3.1. Graph-Based Dimensionality Reduction Method. Recently, graph-based dimensionality reduction (DR) methods become more and more popular in pattern recognition.

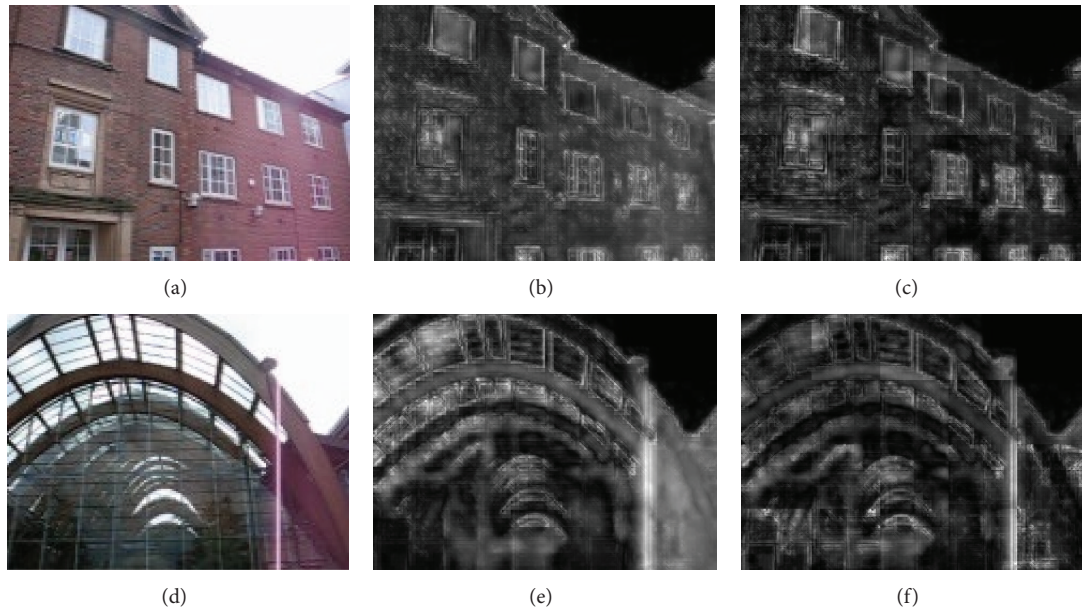


Figure 2: The effect of nonuniform illumination on saliency. (a) and (d) are the original maps selected from the Sheffield buildings database. (b) and (e) are saliency maps calculated by Itti's model. (c) and (f) are subregions based saliency maps.

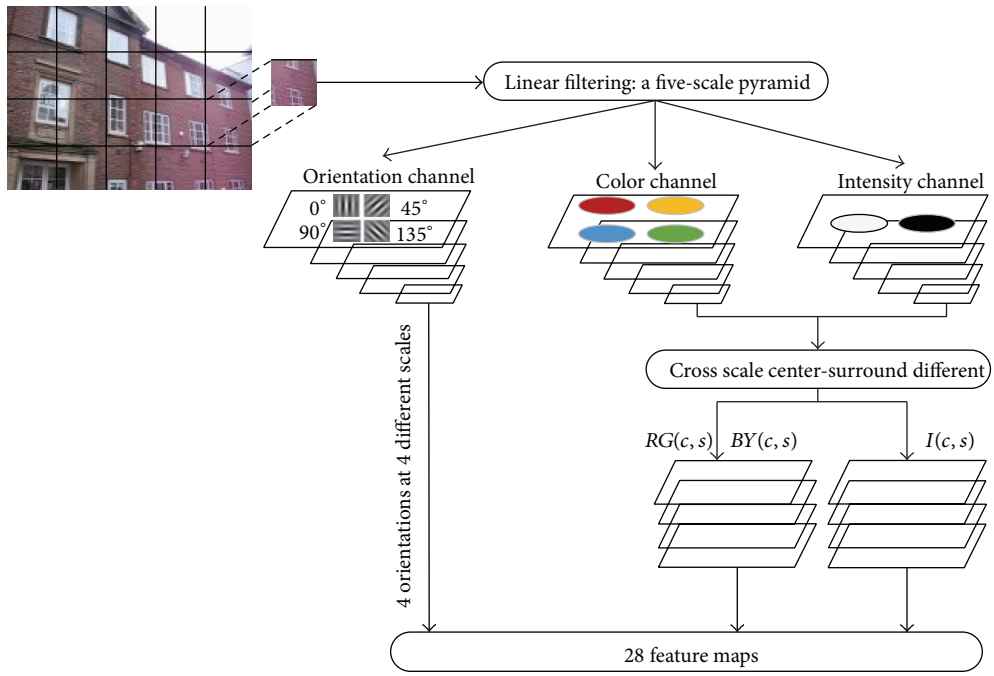


Figure 3: Each subregion's visual feature extract progress.

Yan et al. [27] proposed a graph embedding framework and stated that most of the DR methods [8, 9, 13, 14, 28] could be considered as instances of this framework. Graph construction plays a vital role in improving the performance of graph-based DR algorithms, because graph is a powerful tool, which can capture structural information hidden in the original data. Firstly, we give a review of the traditional graph construction method.

Consider a set of N sample matrices $A = \{A_1, \ldots, A_N\}$, which are taken from an $(n \times m)$-dimensional space. Then, we transform the original sample matrices into their vectorial forms and denote these vectors as $X = \{X_1, \ldots, X_N\}$, $X_i \in R^D$. Let $G = \{V, E, W\}$ denote the weighted graph where V corresponds to the vectors in the vector set X; E denotes the edge set between the sample pairs and matrix W denotes the weight value of the edge between two samples. The construction of the graph G can be presented as a two-stage model. In the first stage, edges between two samples will be constructed by the k nearest neighbor [27] method; namely, samples i and j are connected by an edge if i is one of the k

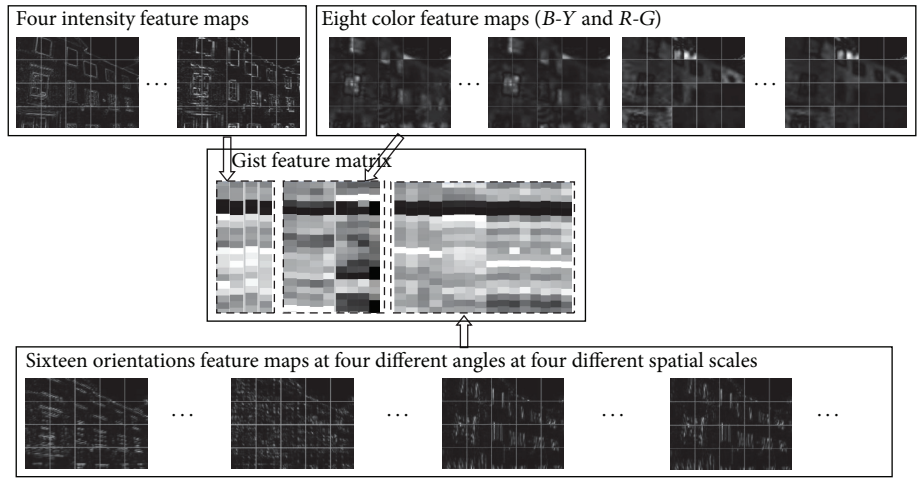


FIGURE 4: The process of building a gist feature matrix. We take the mean of each subregion in each assembled feature map to produce 20 values for a gist column. In total, 28 gist columns are computed. Then all the gist feature columns are combined into a gist feature matrix.

nearest neighbors of j, or j is one of the k nearest neighbors of i.

In the second stage, the weight value for each edge is calculated. The weight value of W_{ij} is calculated as in [27].

(i) Heat kernel:

$$W_{ij} = \begin{cases} \exp\left(\dfrac{-\left\|X_i - X_j\right\|^2}{t}\right), & \text{if } X_i \text{ and } X_j \text{ are neighbors} \\ 0, & \text{otherwise,} \end{cases} \tag{6}$$

(ii) 0-1 way:

$$W_{ij} = \begin{cases} 1, & \text{if } X_i \text{ and } X_j \text{ are neighbors} \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

3.2. Our Proposed Method. In Siagian and Itti's scene classification model [7], gist vectors extracted from different feature maps were combined to form a 544-dimensional feature vector. This representation may ignore some features which are useful to discriminate the original image from others. We note that each column of the feature matrix is a gist feature vector which corresponds to a visual feature of the original building image. Inspired by this, we proposed corresponding columns information based dimensionality reduction (CCI-DR) method which is particularly suitable for dimensional reduction on gist feature matrices. CCI-DR is based on feature matrices' corresponding columns information, namely, the corresponding gist feature vectors, to determine whether two feature matrices can be neighbors when constructing neighbor graph. We name the graph constructed by CCI-DR as the corresponding columns information based graph (CCIG).

Suppose $A = \{A_1, \ldots, A_N\}$ is a building image set taken from an $(n \times m)$-dimensional space. Let l ($l = 1, 2 \cdots m$, m is the maximum column number of the sample matrix) denote the column number of each sample matrix. There are three steps to construct CCIG $= \{V, E, W\}$.

Step 1. Instead of calculating the k nearest neighbors of each sample, we calculate the k nearest neighbors of each column of each sample. For example, let A_i^l denote the lth column of sample A_i. We consider that A_i^l and A_j^l are column neighbors if A_j^l is among the k nearest column neighbors of A_i^l or A_i^l is among the k nearest column neighbors of A_j^l.

Figure 5 shows column A_i^l's three nearest column neighbors. We use ten large rectangles to represent ten sample matrices, and black dots in these rectangles denote the values of matrices. In addition, small bold rectangular boxes represent the column l of each sample.

Step 2. We define sample similarity as the number of column neighbors between two samples. The higher the similarity the two samples have, the more similar the two samples are. S_{ij} is used to denote the number of column neighbors between samples A_i and A_j. When deciding which sample pairs can be neighbors of each other, we only consider the sample pairs whose sample similarity is nonzero. This is because if there is no nearest neighbor column between two sample matrices, the two samples are not similar at all, and consequently they cannot be neighbors. The adjacency matrix E of CCIG is constructed according to the following equation:

$$E_{ij} = \begin{cases} 1, & \text{if } S_{ij} > \dfrac{\left\|S_{i\bullet}\right\|_1}{\left\|S_{i\bullet}\right\|_0} \\ 0, & \text{otherwise,} \end{cases} \tag{8}$$

where $S_{i\bullet} = [S_{ij}]$, $j = 1, \ldots, N-1$, and $S_{i\bullet}$ is a vector which denotes the sample similarity of sample A_i for all other samples. $\|\bullet\|_0$ denotes the L^0-norm, which represents the number of nonzero entries in a vector. So, $\|S_{i\bullet}\|_0$ denotes the number of nonzero entries in vector $S_{i\bullet}$. Let $\|\bullet\|_1$ denote the L^1-norm, which is the linear combination of absolute value of each entry in a vector. So, $\|S_{i\bullet}\|_1$ is the summation of entries in vector $S_{i\bullet}$.

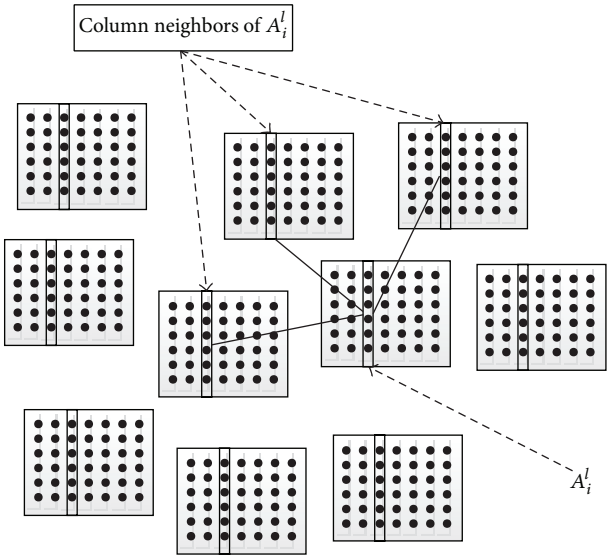


FIGURE 5: The 3 nearest neighbors of sample A_i's lth column.

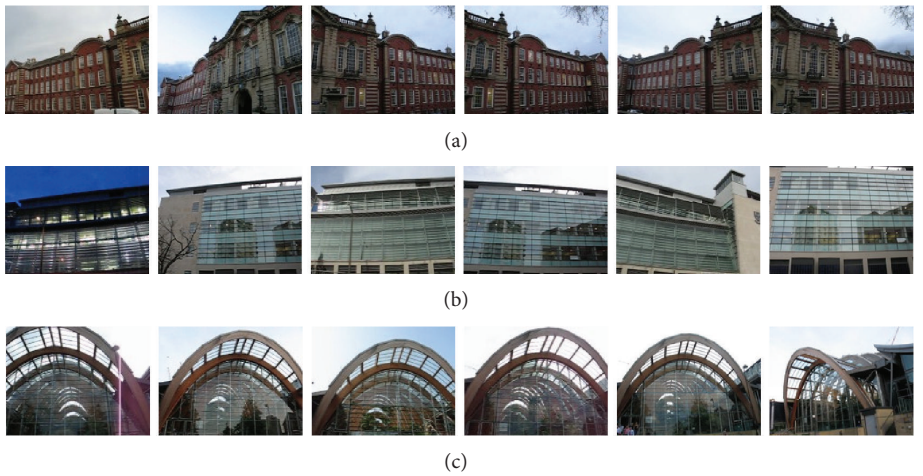


FIGURE 6: (a)–(c) are sample images from categories 1, 10, and 31, respectively.

Formula (8) means that if sample similarity between samples A_i and A_j is greater than the mean of sample similarities between A_i and all other samples, A_j will be considered as a neighbor of A_i, and we put an edge between them; that is, $E_{ij} = 1$.

Step 3. We utilize (6) to compute the weight value for each edge.

The construction method of CCIG is in a general manner, which can be used in most of graph-based dimensionality reduction algorithms. In this research, we incorporate CCIG into a state-of-the-art unsupervised dimensionality reduction algorithm: locality preserving projection (LPP) [9] and develop a new dimensionality reduction algorithm called CCIG-LPP. Similar to LPP, CCIG-LPP aims to preserve gist feature matrices' local manifold structures in high-dimensional space. CCIG-LPP is used to determine whether two feature matrices can be neighbors by comparing corresponding gist feature columns. CCIG-LPP will not mix gist vectors corresponding to different features together, and the features across different building images can be compared. So, CCIG-LPP is particularly suitable for gist feature matrices' dimensional reduction.

4. Experiments

To evaluate the performance of SM-gist and CCI-DR, we carry out a series of experiments on the Sheffield buildings database [29]. The Sheffield buildings database includes 40 buildings, and for each building, the number of building images varies from 100 to 400. There are in total 3192 images with size 160 × 120. Some sample images of the Sheffield building database are shown in Figure 6, and we can see that the database contains many challenging images, because

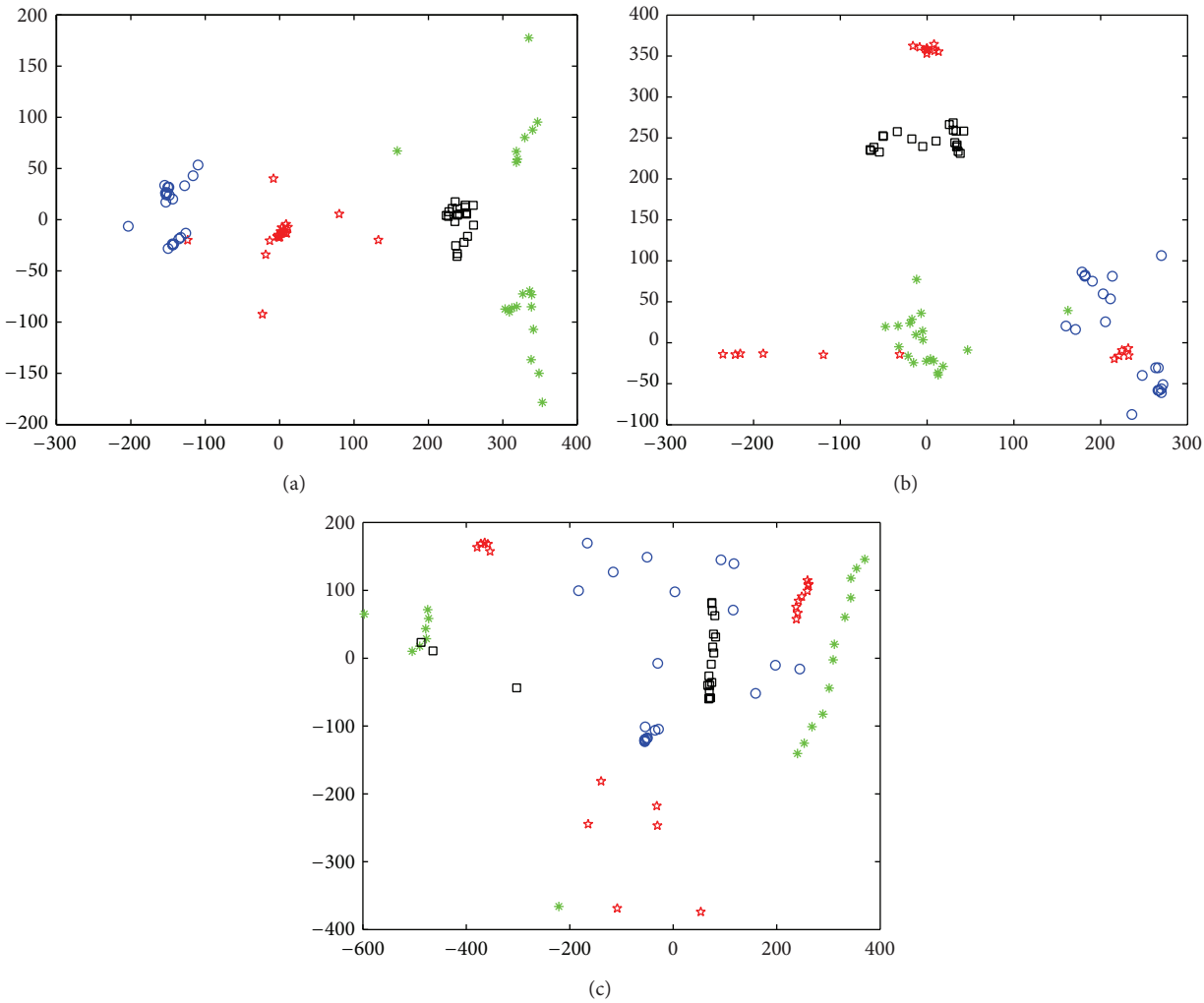


Figure 7: The two-dimensional visualization of the three algorithms. (a) CCIG-LPP, (b) LPP, and (c) PCA.

building images may be taken from various viewpoints, and images may have different scaling and illumination conditions, and there may exist occlusion and rotation phenomena.

The number of building images in each category is different in the original database. We select a subset which we name as $S1$ from the original database. $S1$ consists of 40 categories, and we select the first 20 images for each building.

4.1. Validation of CCI-DR

4.1.1. Gist Feature Manifold Visualization. In this experiment, we randomly selected 4 categories from $S1$, and 20 images form each category. We first used SM-gist model to extract these images' gist features and obtain gist feature matrices of these images. All these gist feature matrices were projected to the 2-dimensional subspace by CCIG-LPP, LPP, and PCA, respectively. Figure 7 shows the 2-dimensional visualization effect of the three algorithms. From Figure 7 we can see the followings.

(1) As shown in Figure 7(a), the samples belonging to four different classes are well separated in the 2-dimensional subspace projected by CCIG-LPP.

(2) Figure 7(b) shows the 2-dimensional subspace projected by LPP, and we can see that the categories represented by red stars were dispersed into three locations, and parts of them were mixed with blue circles.

(3) From Figure 7(c) we can see that, in the subspace projected by PCA, the samples are basically entangled together.

(4) In summary, CCIG-LPP well preserved local intraclass geometry while maximized the local interclass margin separability of different categories. This means that it can maintain more discriminative information for building classification.

4.1.2. Experiment for Accuracies and Different Reduced Dimensions. In this experiment, m ($m = 3$ and 4) images of each category are randomly selected from $S1$ as training samples, and the rest are used for testing. Each experiment was repeated 50 times, and the final results were obtained by taking the average values of the 50 trials. Figure 8 shows the variation of accuracies with different dimensions using LPP

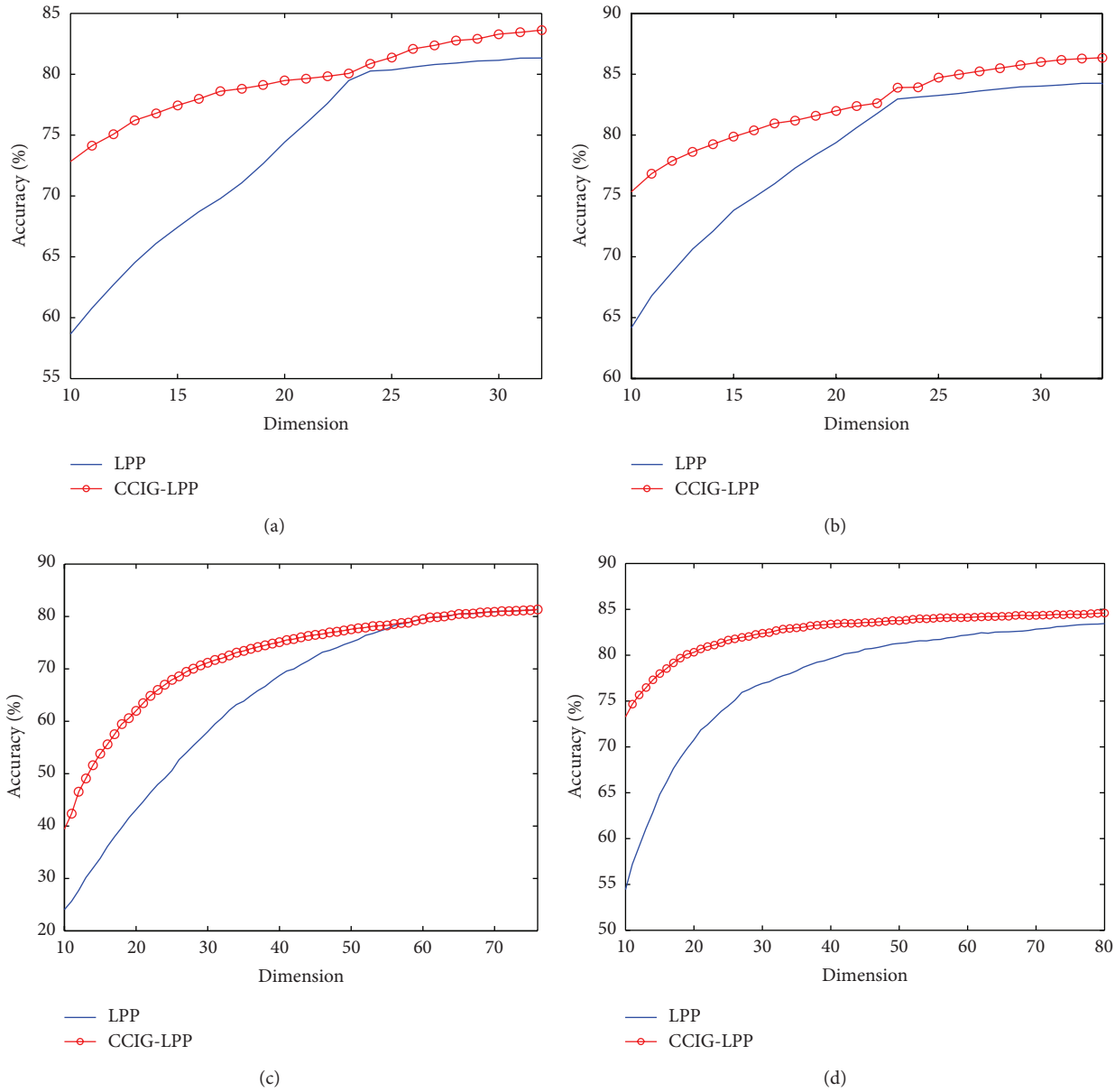


Figure 8: The recognition accuracies of LPP and CCIG-LPP versus the dimensions when three and four images were randomly selected for each category for training. In (a) and (c) three respective images of each category were randomly selected for training. In (b) and (d) four respective images of each category were randomly selected for training. In (a) and (b) SM-gist was used to extract gist features. In (c) and (d) we use Siagian and Itti's feature extraction model to extract gist features.

and CCIG-LPP. Firstly, we use our proposed SM-gist feature extraction model to conduct this experiment. We use LPP and CCIG-LPP on SM-gist feature space for gist feature matrices' dimensional reduction, respectively (as shown in Figures 8(a) and 8(b)). Then, we use Siagian and Itti's feature extraction model to replace SM-gist and repeat this experiment (as shown in Figures 8(c) and 8(d)).

We can see from Figure 8 that CCIG-LPP outperforms LPP in most cases, which is especially obvious in the low dimension situation, regardless which kind of gist feature extraction model is used. This indicates the use of CCIG-LPP as the dimensionality reduction algorithm will make our building recognition model perform even better.

4.2. Building Recognition. Firstly, we evaluate the performance of SM-gist by comparing it with Siagian and Itti's gist feature extract model [7], which is shown by the first and second line of Table 1. We choose locality preserving projection (LPP) as the two gist feature extraction models' dimensional reduction algorithm. Secondly, we evaluate our proposed building recognition method: SM-gist and CCIG-LPP, by comparing it with the above methods.

Let G_m/P_n indicate for each category that we select m images from $S1$ for training and the remaining n ($m+n=20$) images are used for testing. We generate 50 random splits, and the final results are obtained by taking the mean of the

Table 1: Recognition accuracy on the subset of Sheffield buildings database of different gist feature extraction model and different dimensional reduction algorithms. The numbers in parentheses indicate the corresponding feature dimensions which give the best results after dimensionality reduction.

	G_3/P_{17}	G_4/P_{16}	G_5/P_{15}	G_6/P_{14}	G_7/P_{13}	G_8/P_{12}	G_9/P_{11}
Siagian&Itti's model + LPP	76.99 (70)	79.63 (80)	81.25 (85)	84.73 (90)	86.55 (95)	87.13 (97)	89.27 (98)
SM-gist + LPP	79.15 (30)	82.48 (35)	85.13 (35)	87.38 (35)	88.95 (35)	89.76 (35)	90.97 (35)
SM-gist + CCIG-LPP	82.09 (28)	84.69 (30)	86.73 (35)	88.42 (35)	89.96 (35)	90.54 (35)	92.95 (35)

recognition accuracies obtained from these 50 trials. The neighbor parameter k for LPP and CCIG-LPP is set to $G_m - 1$.

From the results presented in Table 1, we can see the following.

(1) With the increase of the number of training samples, the mean accuracies of all the models are improved to some extent.

(2) SM-gist+LPP outperforms Siagian and Itti's model+ LPP for all divisions with G_m/P_n, which is more obvious when the number of training samples is smaller. This is because SM-gist has the ability to adapt to nonuniform illumination, which benefits from the subregion based gist feature extraction method.

(3) The same gist extraction model using different dimensionality reduction methods leads to different results. This indicates that the dimensional reduction algorithm is equally vital for building recognition tasks.

(4) Our proposed model (SM-gist + CCIG-LPP) obtains the best performance for building recognition, which demonstrates the effectiveness and feasibility of our model.

5. Conclusions

In this paper, we proposed a novel building recognition method named subregion's multiscale gist feature (SM-gist) extraction and corresponding columns information based dimensionality reduction (CCI-DR). It is acknowledged that the lighting conditions of an image may be various; however, it is also the fact that the illumination may be more uniform in subareas of the image. So, in our approach building images are divided into subregions and gist features were extracted from each subregion individually. We note that each column of the gist feature matrix is a gist feature vector which corresponds to a visual feature of the original building image. So, we proposed the corresponding columns information based dimensionality reduction (CCI-DR) method which is based on feature matrices' corresponding columns information, namely, the corresponding gist feature vectors, which can determine whether two feature matrices can be neighbors when constructing neighbor graph.

Several experiments are conducted on the Sheffield buildings database. Experimental results show that SM-gist outperformed Siagian and Itti's model and CCI-DR is an effective dimensionality reduction algorithm, which is especially suitable for gist feature matrices' dimensional reduction.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research is supported by (1) the Ph.D. Programs Foundation of Ministry of Education of China under Grant no. 20120061110045 and (2) the Natural Science Foundation of Jilin province of China under Grant no. 201115022; (3) Wei Pang is supported by the partnership fund from dot. rural, RCUK Digital Economy research.

References

[1] M. M. Ullah, A. Pronobis, B. Caputo, J. Luo, P. Jensfelt, and H. I. Christensen, "Towards robust place recognition for robot localization," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '08)*, pp. 530–537, May 2008.

[2] H. Ali, G. Paar, and L. Paletta, "Semantic Indexing for visual recognition of buildings," in *Proceedings of the International Symposium on Mobile Mapping Technology*, pp. 28–31, 2007.

[3] R. Hutchings and W. Mayol-Cuevas, "Building recognition for mobile devices: incorporating positional information with visual features," Tech. Rep. CSTR-06-017, Computer Science, University of Bristol, 2005.

[4] Y. Li and L. G. Shapiro, "Consistent line clusters for building recognition in CBIR," in *Proceedings of the IEEE International Conference on Pattern Recognition*, vol. 3, pp. 952–956, 2002.

[5] W. Zhang and J. Košecká, "Hierarchical building recognition," *Image and Vision Computing*, vol. 25, no. 5, pp. 704–716, 2007.

[6] J. Li and N. M. Allinson, "Subspace learning-based dimensionality reduction in building recognition," *Neurocomputing*, vol. 73, no. 1-3, pp. 324–330, 2009.

[7] C. Siagian and L. Itti, "Rapid biologically-inspired scene classification using features shared with visual attention," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 2, pp. 300–312, 2007.

[8] I. T. Jolliffe, *Principal Component Analysis*, Springer, New York, NY, USA, 1989.

[9] X. F. He and P. Niyogi, "Locality preserving projections," in *Proceedings of the Neural Information Processing Systems (NIPS '04)*, vol. 6, pp. 1059–1071, 2004.

[10] G. J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*, John Wiley & Sons, New York, NY, USA, 2004.

[11] H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional data—with application to face recognition," *Pattern Recognition*, vol. 34, pp. 2067–2070, 2001.

[12] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 228–233, 2001.

[13] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[14] X. He, D. Cai, S. Yan, and H.-J. Zhang, "Neighborhood preserving embedding," in *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV '05)*, vol. 2, pp. 1208–1213, October 2005.

[15] S. Yin, S. Ding, A. Haghani, H. Hao, and P. Zhang, "A comparison study of basic datadriven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process," *Journal of Process Control*, vol. 22, pp. 1567–1581, 2012.

[16] S. Yin, G. Wang, and H. Karimi, "Data-driven design of robust fault detection system for wind turbines," *Mechatronics*, 2013.

[17] S. Yin, S. X. Ding, A. H. A. Sari, and H. Hao, "Data-driven monitoring for stochastic systems and its application on batch process," *International Journal of Systems Science*, vol. 44, no. 7, pp. 1366–1376, 2013.

[18] S. Yin, H. Luo, and S. Ding, "Real-time implementation of fault-tolerant control systems with performance optimization," *IEEE Transactions on Industrial Electronics*, vol. 64, pp. 2402–2411, 2014.

[19] X. Zhao, L. Zhang, P. Shi, and H. R. Karimi, "Robust control of continuous-time systems with state-dependent uncertainties and its application to electronic circuits," *IEEE Transactions on Industrial Electronics*, vol. 61, pp. 4161–4170, 2014.

[20] X. Zhao, H. Liu, and J. Zhang, "Multiple-mode observer design for a class of switched linear systems linear systems," *IEEE Transactions on Automation Science and Engineering*, 2013.

[21] J. B. Tenenbaum, V. De Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[22] K. Q. Weinberger, F. Sha, and L. K. Saul, "Learning a kernel matrix for nonlinear dimensionality reduction," in *Proceedings of the 21st International Conference on Machine Learning (ICML '04)*, p. 106, ACM, July 2004.

[23] B. Yang and S. Chen, "Sample-dependent graph construction with application to dimensionality reduction," *Neurocomputing*, vol. 74, no. 1–3, pp. 301–314, 2010.

[24] A. Oliva and A. Torralba, "Modeling the shape of the scene: a holistic representation of the spatial envelope," *International Journal of Computer Vision*, vol. 42, no. 3, pp. 145–175, 2001.

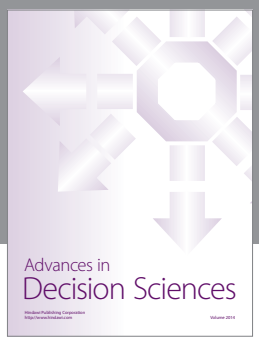
[25] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.

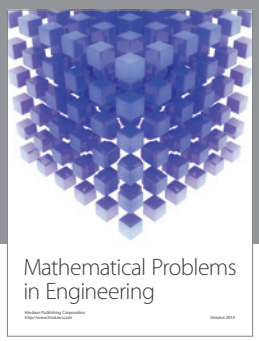
[26] B. Leibe and B. Schiele, "Interleaved object categorization and segmentation," in *Proceedings of the British Machine Vision Conference (BMVC '03)*, pp. 759–768, 2003.

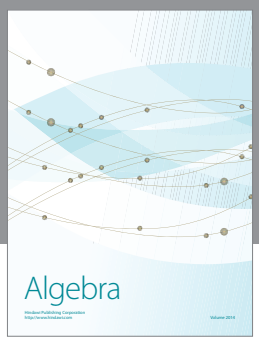
[27] S. Yan, D. Xu, B. Zhang, H.-J. Zhang, Q. Yang, and S. Lin, "Graph embedding and extensions: a general framework for dimensionality reduction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 1, pp. 40–51, 2007.

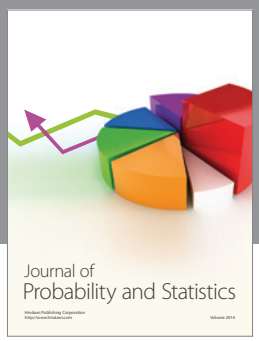
[28] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 586–591, June 1991.

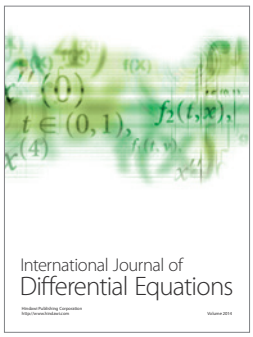
[29] http://eeepro.shef.ac.uk/building/dataset.rar.

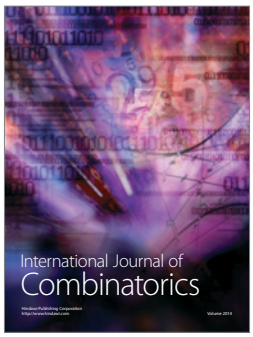


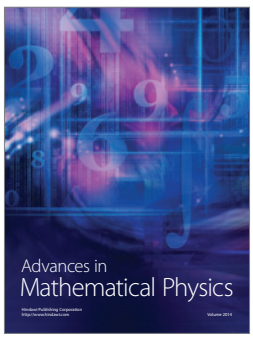


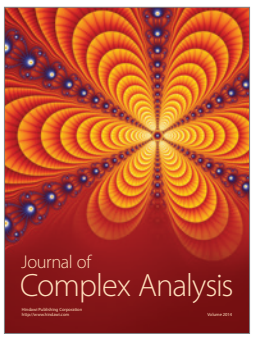


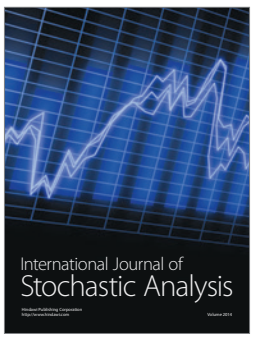


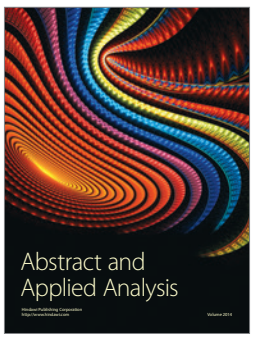


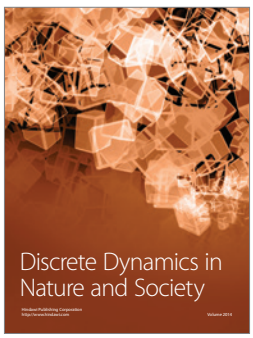


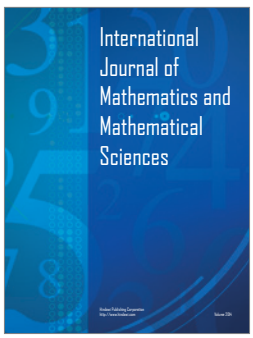


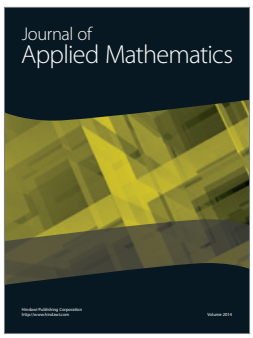












Hindawi
Submit your manuscripts at
http://www.hindawi.com